

AUDIOVISUAL INFORMATION MANAGEMENT SYSTEM

BACKGROUND OF THE INVENTION

5 The present invention relates to a system for managing audiovisual information, and in particular to a system for audiovisual information browsing, filtering, searching, archiving, and personalization.

10 Video cassette recorders (VCRs) may record video programs in response to pressing a record button or may be programmed to record video programs based on the time of day. However, the viewer must program the VCR based on information from a television guide to identify relevant programs to record. After recording, the viewer scans through the entire video tape to select relevant portions of the program for viewing using the
15 functionality provided by the VCR, such as fast forward and fast reverse. Unfortunately, the searching and viewing is based on a linear search, which may require significant time to locate the desired portions of the program(s) and fast forward to the desired portion of the
20 tape. In addition, it is time consuming to program the VCR in light of the television guide to record desired programs. Also, unless the viewer recognizes the programs from the television guide as desirable it is unlikely that the viewer will select such programs to be
25 recorded.

RePlayTV and TiVo have developed hard disk based systems that receive, record, and play television broadcasts in a manner similar to a VCR. The systems may
30 be programmed with the viewer's viewing preferences. The systems use a telephone line interface to receive scheduling information similar to that available from a television guide. Based upon the system programming and the scheduling information, the system automatically
35 records programs that may be of potential interest to the viewer. Unfortunately, viewing the recorded programs occurs in a linear manner and may require substantial

time. In addition, each system must be programmed for an individual's preference, likely in a different manner.

Freeman et al., U.S. Patent No. 5,861,881, disclose an interactive computer system where subscribers
5 can receive individualized content.

With all the aforementioned systems, each individual viewer is required to program the device according to his particular viewing preferences. Unfortunately, each different type of device has
10 different capabilities and limitations which limit the selections of the viewer. In addition, each device includes a different interface which the viewer may be unfamiliar with. Further, if the operator's manual is inadvertently misplaced it may be difficult for the
15 viewer to efficiently program the device.

BRIEF SUMMARY OF THE INVENTION

The present invention overcomes the aforementioned drawbacks of the prior art by providing a method of using a system, which may include, at least one
20 of audio, image, and a video comprising a plurality of frames. A usage preferences description, describing preferences of a user with respect to the use of at least one of the audio, image, and video, where the description normally includes multiple preferences. In one aspect, a
25 protection attribute with respect to at least one of the preferences indicates whether one of the preferences is considered public or private. Other aspects of the user preferences description include other attributes, alone
30 or in combination.

The foregoing and other objectives, features and advantages of the invention will be more readily understood upon consideration of the following detailed description of the invention, taken in conjunction with
35 the accompanying drawings.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

5 FIG. 2 is an exemplary embodiment of the
audiovisual system, including an analysis module, of
FIG. 1.

10 FIG. 4 is an illustration of a thumbnail view
(category) for the audiovisual system.

FIG. 6 is an illustration of a text view
15 (channel) for the audiovisual system.

FIG. 8 is an illustration of a shot view for the audiovisual system.

FIG. 10 is an illustration of a highlight view for the audiovisual system.

FIG. 12 is an illustration of a character/object view for the audiovisual system.

FIG. 14 is an exemplary embodiment of the visualization description scheme of FIG. 13.

35 FIG. 15 is an exemplary embodiment of the meta
information description scheme of FIG. 13.

FIG. 16 is an exemplary embodiment of a segment description scheme for the syntactic structure description scheme of FIG. 13.

5 FIG. 17 is an exemplary embodiment of a region description scheme for the syntactic structure description scheme of FIG. 13.

FIG. 18 is an exemplary embodiment of a segment/region relation description scheme for the syntactic structure description scheme of FIG. 13.

10 FIG. 19 is an exemplary embodiment of an event description scheme for the semantic structure description scheme of FIG. 13.

FIG. 20 is an exemplary embodiment of an object description scheme for the semantic structure description scheme of FIG. 13.

FIG. 21 is an exemplary embodiment of an event/object relation graph description scheme for the syntactic structure description scheme of FIG. 13.

20 FIG. 22 is an exemplary embodiment of a user preference description scheme.

FIG. 23 is an exemplary embodiment of the interrelationship between a usage history description scheme, an agent, and the usage preference description scheme of FIG. 22.

25 FIG. 24 is an exemplary embodiment of the interrelationship between audio and/or video programs together with their descriptors, user identification, and the usage preference description scheme of FIG. 22.

30 FIG. 25 is an exemplary embodiment of a usage preference description scheme of FIG. 22.

FIG. 26 is an exemplary embodiment of the interrelationship between the usage description schemes and an MPEG-7 description schemes.

35 FIG. 27 is an exemplary embodiment of a usage history description scheme of FIG. 22.

FIG. 28 is an exemplary system incorporating the user history description scheme.

002040"80344560

FIG. 29 is an exemplary user preferences description scheme.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

5 Many households today have many sources of audio and video information, such as multiple television sets, multiple VCR's, a home stereo, a home entertainment center, cable television, satellite television, internet broadcasts, world wide web, data services, specialized
10 Internet services, portable radio devices, and a stereo in each of their vehicles. For each of these devices, a different interface is normally used to obtain, select, record, and play the video and/or audio content. For example, a VCR permits the selection of the recording
15 times but the user has to correlate the television guide with the desired recording times. Another example is the user selecting a preferred set of preselected radio stations for his home stereo and also presumably selecting the same set of preselected stations for each
20 of the user's vehicles. If another household member desires a different set of preselected stereo selections, the programming of each audio device would need to be reprogrammed at substantial inconvenience.

The present inventors came to the realization
25 that users of visual information and listeners to audio information, such as for example radio, audio tapes, video tapes, movies, and news, desire to be entertained and informed in more than merely one uniform manner. In other words, the audiovisual information presented to a
30 particular user should be in a format and include content suited to their particular viewing preferences. In addition, the format should be dependent on the content of the particular audiovisual information. The amount of information presented to a user or a listener should be
35 limited to only the amount of detail desired by the particular user at the particular time. For example with the ever increasing demands on the user's time, the user

may desire to watch only 10 minutes of or merely the highlights of a basketball game. In addition, the present inventors came to the realization that the necessity of programming multiple audio and visual devices with their particular viewing preferences is a burdensome task, especially when presented with unfamiliar recording devices when traveling. When traveling, users desire to easily configure unfamiliar devices, such as audiovisual devices in a hotel room, with their viewing and listening preferences in a efficient manner.

The present inventors came to the further realization that a convenient technique of merely recording the desired audio and video information is not sufficient because the presentation of the information should be in a manner that is time efficient, especially in light of the limited time frequently available for the presentation of such information. In addition, the user should be able to access only that portion of all of the available information that the user is interested in, while skipping the remainder of the information.

A user is not capable of watching or otherwise listening to the vast potential amount of information available through all, or even a small portion of, the sources of audio and video information. In addition, with the increasing information potentially available, the user is not likely even aware of the potential content of information that he may be interested in. In light of the vast amount of audio, image, and video information, the present inventors came to the realization that a system that records and presents to the user audio and video information based upon the user's prior viewing and listening habits, preferences, and personal characteristics, generally referred to as user information, is desirable. In addition, the system may present such information based on the capabilities of the system devices. This permits the system to record

desirable information and to customize itself automatically to the user and/or listener. It is to be understood that user, viewer, and/or listener terms may be used interchangeability for any type of content.

5 Also, the user information should be portable between and usable by different devices so that other devices may likewise be configured automatically to the particular user's preferences upon receiving the viewing information.

10 In light of the foregoing realizations and motivations, the present inventors analyzed a typical audio and video presentation environment to determine the significant portions of the typical audiovisual environment. First, referring to FIG. 1 the video,
15 image, and/or audio information 10 is provided or otherwise made available to a user and/or a (device) system. Second, the video, image, and/or audio information is presented to the user from the system 12 (device), such as a television set or a radio. Third,
20 the user interacts both with the system (device) 12 to view the information 10 in a desirable manner and has preferences to define which audio, image, and/or video information is obtained in accordance with the user information 14. After the proper identification of the
25 different major aspects of an audiovisual system the present inventors then realized that information is needed to describe the informational content of each portion of the audiovisual system 16.

30 With three portions of the audiovisual presentation system 16 identified, the functionality of each portion is identified together with its interrelationship to the other portions. To define the necessary interrelationships, a set of description schemes containing data describing each portion is
35 defined. The description schemes include data that is auxiliary to the programs 10, the system 12, and the user 14, to store a set of information, ranging from human

readable text to encoded data, that can be used in enabling browsing, filtering, searching, archiving, and personalization. By providing a separate description scheme describing the program(s) 10, the user 14, and the system 12, the three portions (program, user, and system) may be combined together to provide an interactivity not previously achievable. In addition, different programs 10, different users 14, and different systems 12 may be combined together in any combination, while still maintaining full compatibility and functionality. It is to be understood that the description scheme may contain the data itself or include links to the data, as desired.

A program description scheme 18 related to the video, still image, and/or audio information 10 preferably includes two sets of information, namely, program views and program profiles. The program views define logical structures of the frames of a video that define how the video frames are potentially to be viewed suitable for efficient browsing. For example the program views may contain a set of fields that contain data for the identification of key frames, segment definitions between shots, highlight definitions, video summary definitions, different lengths of highlights, thumbnail set of frames, individual shots or scenes, representative frame of the video, grouping of different events, and a close-up view. The program view descriptions may contain thumbnail, slide, key frame, highlights, and close-up views so that users can filter and search not only at the program level but also within a particular program. The description scheme also enables users to access information in varying detail amounts by supporting, for example, a key frame view as a part of a program view providing multiple levels of summary ranging from coarse to fine. The program profiles define distinctive characteristics of the content of the program, such as actors, stars, rating, director, release date, time stamps, keyword identification, trigger profile, still

5
10
15

20
25
30
35

character profiles may include the audio annotation which is separate from and in addition to the associated audio track of the video.

5 The program description scheme may likewise contain similar information regarding audio (such as radio broadcasts) and images (such as analog or digital photographs or a frame of a video).

10 The user description scheme 20 preferably includes the user's personal preferences, and information regarding the user's viewing history such as for example browsing history, filtering history, searching history, and device setting history. The user's personal preferences includes information regarding particular programs and categorizations of programs that the user
15 prefers to view. The user description scheme may also include personal information about the particular user, such as demographic and geographic information, e.g. zip code and age. The explicit definition of the particular programs or attributes related thereto permits the system
20 16 to select those programs from the information contained within the available program description schemes 18 that may be of interest to the user.

Frequently, the user does not desire to learn to program the device nor desire to explicitly program the device.
25 In addition, the user description scheme 20 may not be sufficiently robust to include explicit definitions describing all desirable programs for a particular user. In such a case, the capability of the user description scheme 20 to adapt to the viewing habits of the user to
30 accommodate different viewing characteristics not explicitly provided for or otherwise difficult to describe is useful. In such a case, the user description scheme 20 may be augmented or any technique can be used to compare the information contained in the user
35 description scheme 20 to the available information contained in the program description scheme 18 to make selections. The user description scheme provides a

technique for holding user preferences ranging from program categories to program views, as well as usage history. User description scheme information is persistent but can be updated by the user or by an intelligent software agent on behalf of the user at any arbitrary time. It may also be disabled by the user, at any time, if the user decides to do so. In addition, the user description scheme is modular and portable so that users can carry or port it from one device to another, such as with a handheld electronic device or smart card or transported over a network connecting multiple devices. When user description scheme is standardized among different manufacturers or products, user preferences become portable. For example, a user can personalize the television receiver in a hotel room permitting users to access information they prefer at any time and anywhere. In a sense, the user description scheme is persistent and timeless based. In addition, selected information within the program description scheme may be encrypted since at least part of the information may be deemed to be private (e.g., demographics). A user description scheme may be associated with an audiovisual program broadcast and compared with a particular user's description scheme of the receiver to readily determine whether or not the program's intended audience profile matches that of the user. It is to be understood that in one of the embodiments of the invention merely the user description scheme is included.

The system description scheme 22 preferably manages the individual programs and other data. The management may include maintaining lists of programs, categories, channels, users, videos, audio, and images. The management may include the capabilities of a device for providing the audio, video, and/or images. Such capabilities may include, for example, screen size, stereo, AC3, DTS, color, black/white, etc. The

management may also include relationships between any one or more of the user, the audio, and the images in relation to one or more of a program description scheme(s) and a user description scheme(s). In a similar manner the management may include relationships between one or more of the program description scheme(s) and user description scheme(s). It is to be understood that in one of the embodiments of the invention merely the system description scheme is included.

10 The descriptors of the program description scheme and the user description scheme should overlap, at least partially, so that potential desirability of the program can be determined by comparing descriptors representative of the same information. For example, the program and user description scheme may include the same set of categories and actors. The program description scheme has no knowledge of the user description scheme, and vice versa, so that each description scheme is not dependant on the other for its existence. It is not necessary for the description schemes to be fully populated. It is also beneficial not to include the program description scheme with the user description scheme because there will likely be thousands of programs with associated description schemes which if combined with the user description scheme would result in a unnecessarily large user description scheme. It is desirable to maintain the user description scheme small so that it is more readily portable. Accordingly, a system including only the program description scheme and the user description scheme would be beneficial.

30 The user description scheme and the system description scheme should include at least partially overlapping fields. With overlapping fields the system can capture the desired information, which would otherwise not be recognized as desirable. The system description scheme preferably includes a list of users and available programs. Based on the master list of

available programs, and associated program description scheme, the system can match the desired programs. It is also beneficial not to include the system description scheme with the user description scheme because there will likely be thousands of programs stored in the system description schemes which if combined with the user description scheme would result in a unnecessarily large user description scheme. It is desirable to maintain the user description scheme small so that it is more readily portable. For example, the user description scheme may include radio station preselected frequencies and/or types of stations, while the system description scheme includes the available stations for radio stations in particular cities. When traveling to a different city the user description scheme together with the system description scheme will permit reprogramming the radio stations. Accordingly, a system including only the system description scheme and the user description scheme would be beneficial.

The program description scheme and the system description scheme should include at least partially overlapping fields. With the overlapping fields, the system description scheme will be capable of storing the information contained within the program description scheme, so that the information is properly indexed. With proper indexing, the system is capable of matching such information with the user information, if available, for obtaining and recording suitable programs. If the program description scheme and the system description scheme were not overlapping then no information would be extracted from the programs and stored. System capabilities specified within the system description scheme of a particular viewing system can be correlated with a program description scheme to determine the views that can be supported by the viewing system. For instance, if the viewing device is not capable of playing back video, its system description scheme may describe

its viewing capabilities as limited to keyframe view and slide view only. Program description scheme of a particular program and system description scheme of the viewing system are utilized to present the appropriate views to the viewing system. Thus, a server of programs serves the appropriate views according to a particular viewing system's capabilities, which may be communicated over a network or communication channel connecting the server with user's viewing device. It is preferred to maintain the program description scheme separate from the system description scheme because the content providers repackage the content and description schemes in different styles, times, and formats. Preferably, the program description scheme is associated with the program, even if displayed at a different time. Accordingly, a system including only the system description scheme and the program description scheme would be beneficial.

By preferably maintaining the independence of each of the three description schemes while having fields that correlate the same information, the programs 10, the users 14, and the system 12 may be interchanged with one another while maintaining the functionality of the entire system 16. Referring to FIG. 2, the audio, visual, or audiovisual program 38, is received by the system 16. The program 38 may originate at any suitable source, such as for example broadcast television, cable television, satellite television, digital television, Internet broadcasts, world wide web, digital video discs, still images, video cameras, laser discs, magnetic media, computer hard drive, video tape, audio tape, data services, radio broadcasts, and microwave communications. The program description stream may originate from any suitable source, such as for example PSIP/DVB-SI information in digital television broadcasts, specialized digital television data services, specialized Internet services, world wide web, data files, data over the

telephone, and memory, such as computer memory. The program, user, and/or system description scheme may be transported over a network (communication channel). For example, the system description scheme may be transported to the source to provide the source with views or other capabilities that the device is capable of using. In response, the source provides the device with image, audio, and/or video content customized or otherwise suitable for the particular device. The system 16 may include any device(s) suitable to receive any one or more of such programs 38. An audiovisual program analysis module 42 performs an analysis of the received programs 38 to extract and provide program related information (descriptors) to the description scheme (DS) generation module 44. The program related information may be extracted from the data stream including the program 38 or obtained from any other source, such as for example data transferred over a telephone line, data already transferred to the system 16 in the past, or data from an associated file. The program related information preferably includes data defining both the program views and the program profiles available for the particular program 38. The analysis module 42 performs an analysis of the programs 38 using information obtained from (i) automatic audio-video analysis methods on the basis of low-level features that are extracted from the program(s), (ii) event detection techniques, (iii) data that is available (or extractable) from data sources or electronic program guides (EPGs, DVB-SI, and PSIP), and (iv) user information obtained from the user description scheme 20 to provide data defining the program description scheme.

The selection of a particular program analysis technique depends on the amount of readily available data and the user preferences. For example, if a user prefers to watch a 5 minute video highlight of a particular program, such as a basketball game, the analysis module

42 may invoke a knowledge based system 90 (FIG. 3) to determine the highlights that form the best 5 minute summary. The knowledge based system 90 may invoke a commercial filter 92 to remove commercials and a slow motion detector 54 to assist in creating the video summary. The analysis module 42 may also invoke other modules to bring information together (e.g., textual information) to author particular program views. For example, if the program 38 is a home video where there is no further information available then the analysis module 42 may create a key-frame summary by identifying key-frames of a multi-level summary and passing the information to be used to generate the program views, and in particular a key frame view, to the description scheme. Referring also to FIG. 3, the analysis module 42 may also include other sub-modules, such as for example, a de-mux/decoder 60, a data and service content analyzer 62, a text processing and text summary generator 64, a close caption analyzer 66, a title frame generator 68, an analysis manager 70, an audiovisual analysis and feature extractor 72, an event detector 74, a key-frame summarizer 76, and a highlight summarizer 78.

The generation module 44 receives the system information 46 for the system description scheme. The system information 46 preferably includes data for the system description scheme 22 generated by the generation module 44. The generation module 44 also receives user information 48 including data for the user description scheme. The user information 48 preferably includes data for the user description scheme generated within the generation module 44. The user input 48 may include, for example, meta information to be included in the program and system description scheme. The user description scheme (or corresponding information) is provided to the analysis module 42 for selective analysis of the program(s) 38. For example, the user description scheme may be suitable for triggering the highlight generation

5

10

30

35

PM and the 20/20 show will start in an hour at 10 PM. The user is interested in the sporting events of the week, and all the news about the Microsoft case with the Department of Justice. The user description scheme may
5 include a profile indicating a desire that the particular user wants to obtain all available information regarding the Microsoft trial and selected sporting events for particular teams. In addition, the system description scheme and program description scheme provide information
10 regarding the content of the available information that may selectively be obtained and recorded. The system, in an autonomous manner, periodically obtains and records the audiovisual information that may be of interest to the user during the past week based on the three
15 description schemes. The device most likely has recorded more than one hour of audiovisual information so the information needs to be condensed in some manner. The user starts interacting with the system with a pointer or voice commands to indicate a desire to view recorded
20 sporting programs. On the display, the user is presented with a list of recorded sporting events including Basketball and Soccer. Apparently the user's favorite Football team did not play that week because it was not recorded. The user is interested in basketball games and
25 indicates a desire to view games. A set of title frames is presented on the display that captures an important moment of each game. The user selects the Chicago Bulls game and indicates a desire to view a 5 minute highlight of the game. The system automatically generates
30 highlights. The highlights may be generated by audio or video analysis, or the program description scheme includes data indicating the frames that are presented for a 5 minute highlight. The system may have also recorded web-based textual information regarding the
35 particular Chicago-Bulls game which may be selected by the user for viewing. If desired, the summarized information may be recorded onto a storage device, such

as a DVD with a label. The stored information may also include an index code so that it can be located at a later time. After viewing the sporting events the user may decide to read the news about the Microsoft trial.

5 It is now 9:50 PM and the user is done viewing the news. In fact, the user has selected to delete all the recorded news items after viewing them. The user then remembers to do one last thing before 10 PM in the evening. The next day, the user desires to watch the VHS tape that he
10 received from his brother that day, containing footage about his brother's new baby girl and his vacation to Peru last summer. The user wants to watch the whole 2-hour tape but he is anxious to see what the baby looks like and also the new stadium built in Lima, which was
15 not there last time he visited Peru. The user plans to take a quick look at a visual summary of the tape, browse, and perhaps watch a few segments for a couple of minutes, before the user takes his daughter to her piano lesson at 10 AM the next morning. The user plugs in the
20 tape into his VCR, that is connected to the system, and invokes the summarization functionality of the system to scan the tape and prepare a summary. The user can then view the summary the next morning to quickly discover the baby's looks, and playback segments between the key-
25 frames of the summary to catch a glimpse of the crying baby. The system may also record the tape content onto the system hard drive (or storage device) so the video summary can be viewed quickly. It is now 10:10 PM, and it seems that the user is 10 minutes late for viewing
30 20/20. Fortunately, the system, based on the three description schemes, has already been recording 20/20 since 10 PM. Now the user can start watching the recorded portion of 20/20 as the recording of 20/20 proceeds. The user will be done viewing 20/20 at 11:10
35 PM.

The average consumer has an ever increasing number of multimedia devices, such as a home audio

system, a car stereo, several home television sets, web browsers, etc. The user currently has to customize each of the devices for optimal viewing and/or listening preferences. By storing the user preferences on a
5 removable storage device, such as a smart card, the user may insert the card including the user preferences into such media devices for automatic customization. This results in the desired programs being automatically recorded on the VCR, and setting of the radio stations
10 for the car stereo and home audio system. In this manner the user only has to specify his preferences at most once, on a single device and subsequently, the descriptors are automatically uploaded into devices by the removable storage device. The user description
15 scheme may also be loaded into other devices using a wired or wireless network connection, e.g. that of a home network. Alternatively, the system can store the user history and create entries in the user description scheme based on the user's audio and video viewing habits. In this
20 manner, the user would never need to program the viewing information to obtain desired information. In a sense, the user descriptor scheme enables modeling of the user by providing a central storage for the user's listening, viewing, browsing preferences, and user's behavior. This
25 enables devices to be quickly personalized, and enables other components, such as intelligent agents, to communicate on the basis of a standardized description format, and to make smart inferences regarding the user's preferences.

30 Many different realizations and applications can be readily derived from FIGS. 2 and 3 by appropriately organizing and utilizing their different parts, or by adding peripherals and extensions as needed. In its most general form, FIG. 2 depicts an audiovisual
35 searching, filtering, browsing, and/or recording appliance that is personalizable. The list of more

specific applications/implementations given below is not exhaustive but covers a range.

The user description scheme is a major enabler for personalizable audiovisual appliances. If the structure (syntax and semantics) of the description schemes is known amongst multiple appliances, the user (user) can carry (or otherwise transfer) the information contained within his user description scheme from one appliance to another, perhaps via a smart card--where these appliances support smart card interface-- in order to personalize them. Personalization can range from device settings, such as display contrast and volume control, to settings of television channels, radio stations, web stations, web sites, geographic information, and demographic information such as age, zip code etc. Appliances that can be personalized may access content from different sources. They may be connected to the web, terrestrial or cable broadcast, etc., and they may also access multiple or different types of single media such as video, music, etc.

For example, one can personalize the car stereo using a smart card plugged out of the home system and plugged into the car stereo system to be able to tune to favorite stations at certain times. As another example, one can also personalize television viewing, for example, by plugging the smart card into a remote control that in turn will autonomously command the television receiving system to present the user information about current and future programs that fits the user's preferences. Different members of the household can instantly personalize the viewing experience by inserting their own smart card into the family remote. In the absence of such a remote, this same type of personalization can be achieved by plugging in the smart card directly to the television system. The remote may likewise control audio systems. In another implementation, the television receiving system holds user description schemes for

5

10

15

35

user to browse, search, and filter audiovisual programs. Browsing in the system offers capabilities that are well beyond fast forwarding and rewinding. For instance, the user can view a thumbnail view of different categories of programs stored in the system. The user then may choose frame view, shot view, key frame view, or highlight view, depending on their availability and user's preference. These views can be readily invoked using the relevant information in the program description scheme, especially in program views. The user at any time can start viewing the program either in parts, or in its entirety.

In this application, the program description scheme may be readily available from many services such as: (i) from broadcast (carried by EPG defined as a part of ATSC-PSIP (ATSC-Program Service Integration Protocol) in USA or DVB-SI (Digital Video Broadcast-Service Information) in Europe); (ii) from specialized data services (in addition to PSIP/DVB-SI); (iii) from specialized web sites; (iv) from the media storage unit containing the audiovisual content (e.g., DVD); (v) from advanced cameras (discussed later), and/or may be generated (i.e., for programs that are being stored) by the analysis module 42 or by user input 48.

Contents of digital still and video cameras can be stored and managed by a system that implements the description schemes, e.g., a system as shown in FIG. 2. Advanced cameras can store a program description scheme, for instance, in addition to the audiovisual content itself. The program description scheme can be generated either in part or in its entirety on the camera itself via an appropriate user input interface (e.g., speech, visual menu drive, etc.). Users can input to the camera the program description scheme information, especially those high-level (or semantic) information that may otherwise be difficult to automatically extract by the system. Some camera settings and parameters (e.g., date and time), as well as quantities computed in the camera

(e.g., color histogram to be included in the color profile), can also be used in generating the program description scheme. Once the camera is connected, the system can browse the camera content, or transfer the camera content and its description scheme to the local storage for future use. It is also possible to update or add information to the description scheme generated in the camera.

The IEEE 1394 and Havi standard specifications enable this type of "audiovisual content" centric communication among devices. The description scheme API's can be used in the context of Havi to browse and/or search the contents of a camera or a DVD which also contain a description scheme associated with their content, i.e., doing more than merely invoking the PLAY API to play back and linearly view the media.

The description schemes may be used in archiving audiovisual programs in a database. The search engine uses the information contained in the program description scheme to retrieve programs on the basis of their content. The program description scheme can also be used in navigating through the contents of the database or the query results. The user description scheme can be used in prioritizing the results of the user query during presentation. It is possible of course to make the program description scheme more comprehensive depending on the nature of the particular application.

The description scheme fulfills the user's desire to have applications that pay attention and are responsive to their viewing and usage habits, preferences, and personal demographics. The proposed user description scheme directly addresses this desire in its selection of fields and interrelationship to other description schemes. Because the description schemes are modular in nature, the user can port his user description

The proposed description schemes can be incorporated into current products similar to those from TiVo and Replay TV in order to extend their entertainment informational value. In particular, the description scheme will enable audiovisual browsing and searching of programs and enable filtering within a particular program by supporting multiple program views such as the highlight view. In addition, the description scheme will handle programs coming from sources other than television broadcasts for which TiVo and Replay TV are not designed to handle. In addition, by standardization of TiVo and Replay TV type of devices, other products may be interconnected to such devices to extend their capabilities, such as devices supporting an MPEG 7 description. MPEG-7 is the Moving Pictures Experts Group - 7, acting to standardize descriptions and description schemes for audiovisual information. The device may also be extended to be personalized by multiple users, as desired.

Because the description scheme is defined, the intelligent software agents can communicate among themselves to make intelligent inferences regarding the user's preferences. In addition, the development and upgrade of intelligent software agents for browsing and filtering applications can be simplified based on the standardized user description scheme.

The description scheme is multi-modal in the following sense that it holds both high level (semantic) and low level features and/or descriptors. For example, the high and low level descriptors are actor name and motion model parameters, respectively. High level descriptors are easily readable by humans while low level descriptors are more easily read by machines and less understandable by humans. The program description scheme can be readily harmonized with existing EPG, PSIP, and

DVB-SI information facilitating search and filtering of broadcast programs. Existing services can be extended in the future by incorporating additional information using the compliant description scheme.

5 For example, one case may include audiovisual programs that are prerecorded on a media such as a digital video disc where the digital video disc also contains a description scheme that has the same syntax and semantics of the description scheme that the FSB
10 module uses. If the FSB module uses a different description scheme, a transcoder (converter) of the description scheme may be employed. The user may want to browse and view the content of the digital video disc. In this case, the user may not need to invoke the
15 analysis module to author a program description. However, the user may want to invoke his or her user description scheme in filtering, searching and browsing the digital video disc content. Other sources of program information may likewise be used in the same manner.

20 It is to be understood that any of the techniques described herein with relation to video are equally applicable to images (such as still image or a frame of a video) and audio (such as radio).

 An example of an audiovisual interface is shown
25 in FIGS. 4-12 which is suitable for the preferred audiovisual description scheme. Referring to FIG. 4, by selecting the thumbnail function as a function of category provides a display with a set of categories on the left hand side. Selecting a particular category,
30 such as news, provides a set of thumbnail views of different programs that are currently available for viewing. In addition, the different programs may also include programs that will be available at a different time for viewing. The thumbnail views are short video
35 segments that provide an indication of the content of the respective actual program that it corresponds with. Referring to FIG. 5, a thumbnail view of available

004040" 00044560

programs in terms of channels may be displayed, if desired. Referring to FIG. 6, a text view of available programs in terms of channels may be displayed, if desired. Referring to FIG. 7, a frame view of particular programs may be displayed, if desired. A representative frame is displayed in the center of the display with a set of representative frames of different programs in the left hand column. The frequency of the number of frames may be selected, as desired. Also a set of frames are displayed on the lower portion of the display representative of different frames during the particular selected program. Referring to FIG. 8, a shot view of particular programs may be displayed, as desired. A representative frame of a shot is displayed in the center of the display with a set of representative frames of different programs in the left hand column. Also a set of shots are displayed on the lower portion of the display representative of different shots (segments of a program, typically sequential in nature) during the particular selected program. Referring to FIG. 9, a key frame view of particular programs may be displayed, as desired. A representative frame is displayed in the center of the display with a set of representative frames of different programs in the left hand column. Also a set of key frame views are displayed on the lower portion of the display representative of different key frame portions during the particular selected program. The number of key frames in each key frame view can be adjusted by selecting the level. Referring to FIG. 10, a highlight view may likewise be displayed, as desired. Referring to FIG. 11, an event view may likewise be displayed, as desired. Referring to FIG. 12, a character/object view may likewise be displayed, as desired.

An example of the description schemes is shown below in XML. The description scheme may be implemented

in any language and include any of the included descriptions (or more), as desired.

The proposed program description scheme includes three major sections for describing a video program. The first section identifies the described program. The second section defines a number of views which may be useful in browsing applications. The third section defines a number of profiles which may be useful in filtering and search applications. Therefore, the overall structure of the proposed description scheme is as follows:

```
<?XML version="1.0">
<!DOCTYPE MPEG-7 SYSTEM "mpeg-7.dtd">
<ProgramIdentity>
15 <ProgramID> ... </ProgramID>
    <ProgramName> ... </ProgramName>
    <SourceLocation> ... </SourceLocation>
</ProgramIdentity>
<ProgramViews>
20 <ThumbnailView> ... </ThumbnailView>
    <SlideView> ... </SlideView>
    <FrameView> ... </FrameView>
    <ShotView> ... </ShotView>
    <KeyFrameView> ... </KeyFrameView>
25 <HighlightView> ... </HighlightView>
    <EventView> ... </EventView>
    <CloseUpView> ... </CloseUpView>
    <AlternateView> ... </AlternateView>
</ProgramViews>
30 <ProgramProfiles>
    <GeneralProfile> ... </GeneralProfile>
    <CategoryProfile> ... </CategoryProfile>
    <DateTimeProfile> ... </DateTimeProfile>
    <KeywordProfile> ... </KeywordProfile>
35 <TriggerProfile> ... </TriggerProfile>
    <StillProfile> ... </StillProfile>
    <EventProfile> ... </EventProfile>
    <CharacterProfile> ... </CharacterProfile>
    <ObjectProfile> ... </ObjectProfile>
40 <ColorProfile> ... </ColorProfile>
    <TextureProfile> ... </TextureProfile>
    <ShapeProfile> ... </ShapeProfile>
    <MotionProfile> ... </MotionProfile>
```

5

- Program ID

The descriptor <ProgramID> contains a number or a string to identify a program.

<ProgramName> program-name </ProgramName>

15

```
<SourceLocation> source-url </SourceLocation>
```

20

25 <ThumbnailView>

30

- **Slide view**

```
<SlideView> frame-id ... </SlideView>
```

- 5 The descriptor <SlideView> specifies a number of frames in a program which may be viewed as snapshots or in a slide show manner.

- **Frame view**

```
10 <FrameView> start-frame-id end-frame-id </FrameView>
```

The descriptor <FrameView> specifies the start and end frames of a program. This is the most basic view of a program and any program has a frame view.

15 • **Shot view**

```
<ShotView>
  <Shot id=""> start-frame-id end-frame-id display-frame-id </Shot>
  <Shot id=""> start-frame-id end-frame-id display-frame-id </Shot>
20   ...
  </ShotView>
```

- 25 The descriptor <ShotView> specifies a number of shots in a program. The <Shot> descriptor defines the start and end frames of a shot. It may also specify a frame to represent the shot.

- **Key-frame view**

```
<KeyFrameView>
30   <KeyFrames level="">
      <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
      <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
      ...
      </KeyFrames>
35   <KeyFrames level="">
```

004040"00044560

```

    <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
    <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
    ...
  </KeyFrames>
5  ...
  </KeyFrameView>

```

The descriptor <KeyFrameView> specifies key frames in a program. The key frames may be organized in a hierarchical manner and the hierarchy is captured by the descriptor <KeyFrames> with a level attribute. The clips which are associated with each key frame are defined by the descriptor <Clip>. Here the display frame in each clip is the corresponding key frame.

15 • Highlight view

```

  <HighlightView>
    <Highlight length="">
      <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
20    <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
      ...
    </Highlight>
    <Highlight length="">
      <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
25    <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
      ...
    </Highlight>
    ...
  </HighlightView>
30

```

The descriptor <HighlightView> specifies clips to form highlights of a program. A program may have different versions of highlights which are tailored into various time length. The clips are grouped into each version of highlight which is specified by the descriptor <Highlight> with a length attribute.

- **Event view**

```

<EventView>
  <Events name="">
5    <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
    <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
    ...
  </Events>
  <Events name="">
10   <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
    <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
    ...
  </Events>
  ...
15 </EventView>

```

The descriptor <EventView> specifies clips which are related to certain events in a program. The clips are grouped into the corresponding events which are specified by the descriptor <Event> with a name attribute.

- **Close-up view**

```

<CloseUpView>
  <Target name="">
25   <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
    <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
    ...
  </Target>
  <Target name="">
30   <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
    <Clip id=""> start-frame-id end-frame-id display-frame-id </Clip>
    ...
  </Target>
  ...
35 </CloseUpView>

```

The descriptor <CloseUpView> specifies clips which may be zoomed in to certain targets in a program. The clips are grouped into the corresponding targets which are specified by the descriptor <Target> with a name attribute.

- **Alternate view**

```

<AlternateView>
  <AlternateSource id=""> source-url </AlternateSource>
5  <AlternateSource id=""> source-url </AlternateSource>
  ...
</AlternateView>

```

10 The descriptor <AlternateView> specifies sources which may be shown as alternate views of a program. Each alternate view is specified by the descriptor <AlternateSource> with an id attribute. The locate of the source may be specified in URL format.

Program Profiles

- 15 • **General profile**

```

<GeneralProfile>
  <Title> title-text </Title>
  <Abstract> abstract-text </Abstract>
20 <Audio> voice-annotation </Audio>
  <Www> web-page-url </Www>
  <ClosedCaption> yes/no </ClosedCaption>
  <Language> language-name </Language>
  <Rating> rating </Rating>
25 <Length> time </Length>
  <Authors> author-name ... </Authors>
  <Producers> producer-name ... </Producers>
  <Directors> director-name ... </Directors>
  <Actors> actor-name ... </Actors>
30 ...
</GeneralProfile>

```

The descriptor <GeneralProfile> describes the general aspects of a program.

- 35 • **Category profile**

```

<CategoryProfile> category-name ... </CategoryProfile>

```

The descriptor `<CategoryProfile>` specifies the categories under which a program may be classified.

- **Date-time profile**

5
`<DateTimeProfile>`
 `<ProductionDate> date </ProductionDate>`
 `<ReleaseDate> date </ReleaseDate>`
 `<RecordingDate> date </RecordingDate>`
 10 `<RecordingTime> time </RecordingTime>`
 `...`
 `</DateTimeProfile>`

15 The descriptor `<DateTimeProfile>` specifies various date and time information of a program.

- **Keyword profile**

`<KeywordProfile> keyword ... </KeywordProfile>`
 20 The descriptor `<KeywordProfile>` specifies a number of keywords which may be used to filter or search a program.

- **Trigger profile**

`<TriggerProfile> trigger-frame-id ... </TriggerProfile>`
 25 The descriptor `<TriggerProfile>` specifies a number of frames in a program which may be used to trigger certain actions while the playback of the program.

- **Still profile**

30
`<StillProfile>`
 `<Still id="">`
 `<HotRegion id="">`

004010"00811560

```

    <Location> x1 y1 x2 y2 </Location>
    <Text> text-annotation </Text>
    <Audio> voice-annotation </Audio>
    <Www> web-page-url </Www>
5    </HotRegion>
    <HotRegion id ="">
        <Location> x1 y1 x2 y2 </Location>
        <Text> text-annotation </Text>
        <Audio> voice-annotation </Audio>
10    <Www> web-page-url </Www>
    </HotRegion>
    ...
</Still>
<Still id="">
15    <HotRegion id ="">
        <Location> x1 y1 x2 y2 </Location>
        <Text> text-annotation </Text>
        <Audio> voice-annotation </Audio>
        <Www> web-page-url </Www>
20    </HotRegion>
    <HotRegion id ="">
        <Location> x1 y1 x2 y2 </Location>
        <Text> text-annotation </Text>
        <Audio> voice-annotation </Audio>
25    <Www> web-page-url </Www>
    </HotRegion>
    ...
</Still>
    ...
30 </StillProfile>

```

The descriptor <StillProfile> specifies hot regions or regions of interest within a frame. The frame is specified by the descriptor <Still> with an id attribute which corresponds to the frame-id. Within a frame, each hot region is specified by the descriptor <HotRegion> with an id attribute.

- **Event profile**

```

40 <EventProfile>
    <EventList> event-name ... </EventList>
    <Event name="">

```

```

<Www> web-page-url </Www>
<Occurrence id="">
    <Duration> start-frame-id end-frame-id </Duration>
    <Text> text-annotation </Text>
5    <Audio> voice-annotation </Audio>
</Occurrence>
<Occurrence id="">
    <Duration> start-frame-id end-frame-id </Duration>
    <Text> text-annotation </Text>
10    <Audio> voice-annotation </Audio>
</Occurrence>
...
</Event>
<Event name="">
15    <Www> web-page-url </Www>
    <Occurrence id="">
        <Duration> start-frame-id end-frame-id </Duration>
        <Text> text-annotation </Text>
        <Audio> voice-annotation </Audio>
20    </Occurrence>
    <Occurrence id="">
        <Duration> start-frame-id end-frame-id </Duration>
        <Text> text-annotation </Text>
        <Audio> voice-annotation </Audio>
25    </Occurrence>
    ...
</Event>
...
</EventProfile>
30

```

The descriptor <EventProfile> specifies the detailed information for certain events in a program. Each event is specified by the descriptor <Event> with a name attribute. Each occurrence of an event is specified by the descriptor <Occurrence> with an id attribute which

35 may be matched with a clip id under <EventView>.

• Character profile

```

<CharacterProfile>
40    <CharacterList> character-name ... </CharacterList>
    <Character name="">
        <ActorName> actor-name </ActorName>

```

```

<Gender> male </Gender>
<Age> age </Age>
<Www> web-page-url </Www>
<Occurrence id="">
5      <Duration> start-frame-id end-frame-id </Duration>
      <Location> frame:[x1 y1 x2 y2] ... </Location>
      <Motion> vx vy vz vα vβ vγ </Motion>
      <Text> text-annotation </Text>
      <Audio> voice-annotation </Audio>
10    </Occurrence>
    <Occurrence id="">
      <Duration> start-frame-id end-frame-id </Duration>
      <Location> frame:[x1 y1 x2 y2] ... </Location>
      <Motion> vx vy vz vα vβ vγ </Motion>
15    <Text> text-annotation </Text>
      <Audio> voice-annotation </Audio>
    </Occurrence>
    ...
</Character>
20    <Character name="">
      <ActorName> actor-name </ActorName>
      <Gender> male </Gender>
      <Age> age </Age>
      <Www> web-page-url </Www>
25    <Occurrence id="">
      <Duration> start-frame-id end-frame-id </Duration>
      <Location> frame:[x1 y1 x2 y2] ... </Location>
      <Motion> vx vy vz vα vβ vγ </Motion>
      <Text> text-annotation </Text>
30    <Audio> voice-annotation </Audio>
    </Occurrence>
    <Occurrence id="">
      <Duration> start-frame-id end-frame-id </Duration>
      <Location> frame:[x1 y1 x2 y2] ... </Location>
35    <Motion> vx vy vz vα vβ vγ </Motion>
      <Text> text-annotation </Text>
      <Audio> voice-annotation </Audio>
    </Occurrence>
    ...
40    </Character>
    ...
</CharacterProfile>

```

45 The descriptor <CharacterProfile> specifies the detailed information for certain characters in a program. Each character is specified by the descriptor <Character> with

a name attribute. Each occurrence of a character is specified by the descriptor <Occurrence> with an id attribute which may be matched with a clip id under <CloseUpView>.

5 • Object profile

```

<ObjectProfile>
  <ObjectList> object-name ... </ObjectList>
  <Object name="">
10    <Www> web-page-url </Www>
      <Occurrence id="">
          <Duration> start-frame-id end-frame-id </Duration>
          <Location> frame:[x1 y1 x2 y2] ... </Location>
          <Motion> vx vy vz va vb vv </Motion>
15    <Text> text-annotation </Text>
          <Audio> voice-annotation </Audio>
      </Occurrence>
      <Occurrence id="">
          <Duration> start-frame-id end-frame-id </Duration>
20    <Location> frame:[x1 y1 x2 y2] ... </Location>
          <Motion> vx vy vz va vb vv </Motion>
          <Text> text-annotation </Text>
          <Audio> voice-annotation </Audio>
      </Occurrence>
25    ...
  </Object>
  <Object name="">
      <Www> web-page-url </Www>
      <Occurrence id="">
30    <Duration> start-frame-id end-frame-id </Duration>
          <Location> frame:[x1 y1 x2 y2] ... </Location>
          <Motion> vx vy vz va vb vv </Motion>
          <Text> text-annotation </Text>
          <Audio> voice-annotation </Audio>
35    </Occurrence>
      <Occurrence id="">
          <Duration> start-frame-id end-frame-id </Duration>
          <Location> frame:[x1 y1 x2 y2] ... </Location>
          <Motion> vx vy vz va vb vv </Motion>
40    <Text> text-annotation </Text>
          <Audio> voice-annotation </Audio>
      </Occurrence>
      ...

```

004040"00644560

```

    </Object>
    ...
</ObjectProfile>

```

5 The descriptor <ObjectProfile> specifies the detailed
information for certain objects in a program. Each object
is specified by the descriptor <Object> with a name
attribute. Each occurrence of a object is specified by
the descriptor <Occurrence> with an id attribute which
10 may be matched with a clip id under <CloseUpView>.

- **Color profile**

```

    <ColorProfile>
    ...
15 </ColorProfile>

```

The descriptor <ColorProfile> specifies the detailed
color information of a program. All MPEG-7 color
descriptors may be placed under here.

20 • **Texture profile**

```

    <TextureProfile>
    ...
    </TextureProfile>

```

25 The descriptor <TextureProfile> specifies the detailed
texture information of a program. All MPEG-7 texture
descriptors may be placed under here.

- **Shape profile**

```

30
    <ShapeProfile>
    ...
    </ShapeProfile>

```

002000" 80811360

The descriptor <ShapeProfile> specifies the detailed shape information of a program. All MPEG-7 shape descriptors may be placed under here.

- **Motion profile**

5

```
<MotionProfile>
```

```
...
```

```
</MotionProfile>
```

10 The descriptor <MotionProfile> specifies the detailed motion information of a program. All MPEG-7 motion descriptors may be placed under here.

User Description Scheme

The proposed user description scheme includes three major sections for describing a user. The first section identifies the described user. The second section records a number of settings which may be preferred by the user. The third section records some statistics which may reflect certain usage patterns of the user. Therefore, the overall structure of the proposed description scheme is as follows:

```
<?XML version="1.0">
<!DOCTYPE MPEG-7 SYSTEM "mpeg-7.dtd">
<UserIdentity>
  <UserID> ... </UserID>
25 <UserName> ... </UserName>
</UserIdentity>
<UserPreferences>
  <BrowsingPreferences> ... </BrowsingPreferences>
  <FilteringPreferences> ... </FilteringPreferences>
30 <SearchPreferences> ... </SearchPreferences>
  <DevicePreferences> ... </DevicePreferences>
</UserPreferences>
<UserHistory>
  <BrowsingHistory> ... </BrowsingHistory>
35 <FilteringHistory> ... </FilteringHistory>
  <SearchHistory> ... </SearchHistory>
```

002040" 80844560

10 • User ID

10 • User ID

15 The descriptor <UserID> contains a number or a string to
 identify a user.

```
<UserName> user-name </UserName>
```

20 The descriptor <UserName> specifies the name of a user.

20 The descriptor <UserName> specifies the name of a user.

- Browsing preferences

```

25      <BrowsingPreferences>
          <Views>
              <ViewCategory id=""> view-id ... </ViewCategory>
              <ViewCategory id=""> view-id ... </ViewCategory>
              ...
          </Views>
30      <FrameFrequency> frequency ...<FrameFrequency>
          <ShotFrequency> frequency ...<ShotFrequency>
          <KeyFrameLevel> level-id ...<KeyFrameLevel>
          <HighlightLength> length ...<HighlightLength>
          ...
35  </BrowsingPreferences>

```

The descriptor `<BrowsingPreferences>` specifies the browsing preferences of a user. The user's preferred views are specified by the descriptor `<Views>`. For each category, the preferred views are specified by the descriptor `<ViewCategory>` with an `id` attribute which corresponds to the category `id`. The descriptor `<FrameFrequency>` specifies at what interval the frames should be displayed on a browsing slider under the frame view. The descriptor `<ShotFrequency>` specifies at what interval the shots should be displayed on a browsing slider under the shot view. The descriptor `<KeyFrameLevel>` specifies at what level the key frames should be displayed on a browsing slider under the key frame view. The descriptor `<HighlightLength>` specifies which version of the highlight should be shown under the highlight view.

• Filtering preferences

```

<FilteringPreferences>
20   <Categories> category-name ... </Categories>
      <Channels> channel-number ... </Channels>
      <Ratings> rating-id ... </Ratings>
      <Shows> show-name ... </Shows>
      <Authors> author-name ... </Authors>
25   <Producers> producer-name ... </Producers>
      <Directors> director-name ... </Directors>
      <Actors> actor-name ... </Actors>
      <Keywords> keyword ... </Keywords>
      <Titles> title-text ... </Titles>
30   ...
</FilteringPreferences>

```

The descriptor `<FilteringPreferences>` specifies the filtering related preferences of a user.

35 • Search preferences

```

<SearchPreferences>

```

```

    <Categories> category-name ... </Categories>
    <Channels> channel-number ... </Channels>
    <Ratings> rating-id ... </Ratings>
    <Shows> show-name ... </Shows>
5    <Authors> author-name ... </Authors>
    <Producers> producer-name ... </Producers>
    <Directors> director-name ... </Directors>
    <Actors> actor-name ... </Actors>
    <Keywords> keyword ... </Keywords>
10   <Titles> title-text ... </Titles>
    ...
    </SearchPreferences>

```

15 The descriptor <SearchPreferences> specifies the search related preferences of a user.

• Device preferences

```

    <DevicePreferences>
        <Brightness> brightness-value </Brightness>
20    <Contrast> contrast-value </Contrast>
        <Volume> volume-value </Volume>
    </DevicePreferences>

```

25 The descriptor <DevicePreferences> specifies the device preferences of a user.

Usage History

• Browsing history

```

    <BrowsingHistory>
30    <Views>
        <ViewCategory id=""> view-id ... </ViewCategory>
        <ViewCategory id=""> view-id ... </ViewCategory>
        ...
    </Views>
35    <FrameFrequency> frequency ...<FrameFrequency>
    <ShotFrequency> frequency ...<ShotFrequency>
    <KeyFrameLevel> level-id ...<KeyFrameLevel>
    <HighlightLength> length ...<HighlightLength>

```

002040" 80811360

...
</BrowsingHistory>

The descriptor <BrowsingHistory> captures the history of
5 a user's browsing related activities.

- **Filtering history**

```

<FilteringHistory>
  <Categories> category-name ... </Categories>
10  <Channels> channel-number ... </Channels>
    <Ratings> rating-id ... </Ratings>
    <Shows> show-name ... </Shows>
    <Authors> author-name ... </Authors>
    <Producers> producer-name ... </Producers>
15  <Directors> director-name ... </Directors>
    <Actors> actor-name ... </Actors>
    <Keywords> keyword ... </Keywords>
    <Titles> title-text ... </Titles>
    ...
20 </FilteringHistory>

```

The descriptor <FilteringHistory> captures the history of
a user's filtering related activities.

- **Search history**

```

25 <SearchHistory>
    <Categories> category-name ... </Categories>
    <Channels> channel-number ... </Channels>
    <Ratings> rating-id ... </Ratings>
30  <Shows> show-name ... </Shows>
    <Authors> author-name ... </Authors>
    <Producers> producer-name ... </Producers>
    <Directors> director-name ... </Directors>
    <Actors> actor-name ... </Actors>
35  <Keywords> keyword ... </Keywords>
    <Titles> title-text ... </Titles>
    ...
</SearchHistory>

```


System Description Scheme

The proposed system description scheme includes four major sections for describing a user. The first section identifies the described system. The second section keeps a list of all known users. The third section keeps lists of available programs. The fourth section describes the capabilities of the system. Therefore, the overall structure of the proposed description scheme is as follows:

```
<?XML version="1.0">
10 DOCTYPE MPEG-7 SYSTEM "mpeg-7.dtd">
<SystemIdentity>
  <SystemID> ... </SystemID>
  <SystemName> ... </SystemName>
  <SystemSerialNumber> ... </SystemSerialNumber>
15 </SystemIdentity>
<SystemUsers>
  <Users> ... </Users>
</SystemUsers>
<SystemPrograms>
20 <Categories> ... </Categories>
  <Channels> ... </Channels>
  <Programs> ... </Programs>
</SystemPrograms>
<SystemCapabilities>
25 <Views> ... </Views>
</SystemCapabilities>
```

System Identity

• System ID

30

```
<SystemID> system-id </SystemID>
```

The descriptor <SystemID> contains a number or a string to identify a video system or device.

35

• System name

<SystemName> system-name </SystemName>

The descriptor <SystemName> specifies the name of a video system or device.

5 • System serial number

<SystemSerialNumber> system-serial-number </SystemSerialNumber>

10 The descriptor <SystemSerialNumber> specifies the serial number of a video system or device.

System Users

• Users

15 <Users>
 <User>
 <UserID> user-id </UserID>
 <UserName> user-name </UserName>
 </User>
 <User>
 20 <UserID> user-id </UserID>
 <UserName> user-name </UserName>
 </User>
 ...
 </Users>
 25

25 The descriptor <SystemUsers> lists a number of users who have registered on a video system or device. Each user is specified by the descriptor <User>. The descriptor <UserID> specifies a number or a string which should
 30 match with the number or string specified in <UserID> in one of the user description schemes.

002010" 00011560

Programs in the System

• Categories

```

<Categories>
5   <Category>
      <CategoryID> category-id </CategoryID>
      <CategoryName> category-name </CategoryName>
      <SubCategories> sub-category-id ... </SubCategories>
    </Category>
10  <Category>
      <CategoryID> category-id </CategoryID>
      <CategoryName> category-name </CategoryName>
      <SubCategories> sub-category-id ... </SubCategories>
    </Category>
15  ...
    </Categories>

```

The descriptor <Categories> lists a number of categories which have been registered on a video system or device. Each category is specified by the descriptor <Category>. The major-sub relationship between categories is captured by the descriptor < SubCategories>.

• Channels

```

25 <Channels>
    <Channel>
      <ChannelID> channel-id </ChannelID>
      <ChannelName> channel-name </ChannelName>
      <SubChannels> sub-channel-id ... </SubChannels>
30 </Channel>
    <Channel>
      <ChannelID> channel-id </ChannelID>
      <ChannelName> channel-name </ChannelName>
      <SubChannels> sub-channel-id ... </SubChannels>
35 </Channel>
    ...
  </Channels>

```

The descriptor <Channels> lists a number of channels which have been registered on a video system or device.

Each channel is specified by the descriptor <Channel>. The major-sub relationship between channels is captured by the descriptor < SubChannels>.

- **Programs**

```

5      <Programs>
      <CategoryPrograms>
        <CategoryID> category-id </CategoryID>
        <Programs> program-id ... </Programs>
10     </CategoryPrograms>
      <CategoryPrograms>
        <CategoryID> category-id </CategoryID>
        <Programs> program-id ... </Programs>
      </CategoryPrograms>
15     ...
      <ChannelPrograms>
        <ChannelID> channel-id </ChannelID>
        <Programs> program-id ... </Programs>
      </ChannelPrograms>
20     <ChannelPrograms>
        <ChannelID> channel-id </ChannelID>
        <Programs> program-id ... </Programs>
      </ChannelPrograms>
      ...
25 </Programs>

```

The descriptor <Programs> lists programs who are available on a video system or device. The programs are grouped under corresponding categories or channels. Each group of programs are specified by the descriptor <CategoryPrograms> or <ChannelPrograms>. Each program id contained in the descriptor <Programs> should match with the number or string specified in <ProgramID> in one of the program description schemes.

- Views

15

20

25

30

allowing access to the necessary data, preferably in a one-step process. The visualization description scheme 406 provides for several different presentations of the video content (or audio), such as for example, a thumbnail view description scheme 410, a key frame view description scheme 412, a highlight view description scheme 414, an event view description scheme 416, a close-up view description scheme 418, and an alternative view description scheme 420. Other presentation techniques and description schemes may be added, as desired. The thumbnail view description scheme 410 preferably includes an image 422 or reference to an image representative of the video content and a time reference 424 to the video. The key frame view description scheme 412 preferably includes a level indicator 426 and a time reference 428. The level indicator 426 accommodates the presentation of a different number of key frames for the same video portion depending on the user's preference. The highlight view description scheme 414 includes a length indicator 430 and a time reference 432. The length indicator 430 accommodates the presentation of a different highlight duration of a video depending on the user's preference. The event view description scheme 416 preferably includes an event indicator 434 for the selection of the desired event and a time reference 436. The close-up view description scheme 418 preferably includes a target indicator 438 and a time reference 440. The alternate view description scheme preferably includes a source indicator 442. To increase performance of the system it is preferred to specify the data which is needed to render such views in a centralized and straightforward manner. By doing so, it is then feasible to access the data in a simple one-step process without complex parsing of the video.

Referring to FIG. 15, the meta information description scheme 408 generally includes various descriptors which carry general information about a video

(or audio) program such as the title, category, keywords, etc. Additional descriptors, such as those previously described, may be included, as desired.

Referring again to FIG. 13, the syntactic structure description scheme 402 specifies the physical structure of a video program (or audio), e.g., a table of contents. The physical features, may include for example, color, texture, motion, etc. The syntactic structure description scheme 402 preferably includes three modules, namely a segment description scheme 450, a region description scheme 452, and a segment/region relation graph description scheme 454. The segment description scheme 450 may be used to define relationships between different portions of the video consisting of multiple frames of the video. A segment description scheme 450 may contain another segment description scheme 450 and/or shot description scheme to form a segment tree. Such a segment tree may be used to define a temporal structure of a video program. Multiple segment trees may be created and thereby create multiple table of contents. For example, a video program may be segmented into story units, scenes, and shots, from which the segment description scheme 450 may contain such information as a table of contents. The shot description scheme may contain a number of key frame description schemes, a mosaic description scheme(s), a camera motion description scheme(s), etc. The key frame description scheme may contain a still image description scheme which may in turn contains color and texture descriptors. It is noted that various low level descriptors may be included in the still image description scheme under the segment description scheme. Also, the visual descriptors may be included in the region description scheme which is not necessarily under a still image description scheme. On example of a segment description scheme 450 is shown in FIG. 16.

Referring to FIG. 17, the region description scheme 452 defines the interrelationships between groups of pixels of the same and/or different frames of the video. The region description scheme 452 may also
5 contain geometrical features, color, texture features, motion features, etc.

Referring to FIG. 18, the segment/region relation graph description scheme 454 defines the interrelationships between a plurality of regions (or
10 region description schemes), a plurality of segments (or segment description schemes), and/or a plurality of regions (or description schemes) and segments (or description schemes).

Referring again to FIG. 13, the semantic
15 structure description scheme 404 is used to specify semantic features of a video program (or audio), e.g. semantic events. In a similar manner to the syntactic structure description scheme, the semantic structure description scheme 404 preferably includes three modules,
20 namely an event description scheme 480, an object description scheme 482, and an event/object relation graph description scheme 484. The event description scheme 480 may be used to form relationships between different events of the video normally consisting of
25 multiple frames of the video. An event description scheme 480 may contain another event description scheme 480 to form a segment tree. Such an event segment tree may be used to define a semantic index table for a video program. Multiple event trees may be created and thereby
30 creating multiple index tables. For example, a video program may include multiple events, such as a basketball dunk, a fast break, and a free throw, and the event description scheme may contain such information as an index table. The event description scheme may also
35 contain references which link the event to the corresponding segments and/or regions specified in the

syntactic structure description scheme. On example of an event description scheme is shown in FIG. 19.

Referring to FIG. 20, the object description scheme 482 defines the interrelationships between groups of pixels of the same and/or different frames of the video representative of objects. The object description scheme 482 may contain another object description scheme and thereby form an object tree. Such an object tree may be used to define an object index table for a video program. The object description scheme may also contain references which link the object to the corresponding segments and/or regions specified in the syntactic structure description scheme.

Referring to FIG. 21, the event/object relation graph description scheme 484 defines the interrelationships between a plurality of events (or event description schemes), a plurality of objects (or object description schemes), and/or a plurality of events (or description schemes) and objects (or description schemes).

After further consideration, the present inventors came the realization that the particular design of the user preference description scheme is important to implement portability, while permitting adaptive updating, of the user preference description scheme. Moreover, the user preference description scheme should be readily usable by the system while likewise being suitable for modification based on the user's historical usage patterns. It is possible to collectively track all users of a particular device to build a database for the historical viewing preferences of the users of the device, and thereafter process the data dynamically to determine which content the users would likely desire. However, this implementation would require the storage of a large amount of data and the associated dynamic processing requirements to determine the user preferences. It is to be understood that the user

preference description scheme may be used alone or in combination with other description scheme.

Referring to FIG. 22, to achieve portability and potentially decreased processing requirements the user preference description scheme 20 should be divided into at least two separate description schemes, namely, a usage preference description scheme 500 and a usage history description scheme 502. The usage preference description scheme 500, described in detail later, includes a description scheme of the user's audio and/or video consumption preferences. The usage preference description scheme 500 describes one or more of the following, depending on the particular implementation, (a) browsing preferences, (b) filtering preferences, (c) searching preferences, and (d) device preferences of the user. The type of preferences shown in the usage preference description scheme 500 are generally immediately usable by the system for selecting and otherwise using the available audio and/or video content. In other words, the usage preference description scheme 500 includes data describing audio and/or video consumption of the user. The usage history description scheme 502, described in detail later, includes a description scheme of the user's historical audio and/or video activity, such as browsing, device settings, viewing, and selection. The usage history description scheme 502 describes one or more of the following, depending on the particular implementation, (a) browsing history, (b) filtering history, (c) searching history, and (d) device usage history. The type of preferences shown in the usage history description scheme 502 are not generally immediately usable by the system for selecting and otherwise using the available audio and/or video content. The data contained in the usage history description scheme 502 may be considered generally "unprocessed", at least in comparison to the data contained in the usage preferences description scheme 500

because it generally contains the historical usage data of the audio and/or video content of the viewer.

In general, capturing the user's usage history facilitates "automatic" composition of user preferences by a machine, as desired. When updating the user preference description scheme 500 it is desirable that the usage history description scheme 502 be relatively symmetric to the usage preference description scheme 500. The symmetry permits more effective updating because less interpretation between the two description schemes is necessary in order to determine what data should be included in the preferences. Numerous algorithms can then be applied in utilization of the history information in deriving user preferences. For instance, statistics can be computed from the history and utilized for this purpose.

After consideration of the usage preference description 500 and the usage history description 502, the present inventors came to the realization that in the home environment many different users with different viewing and usage preferences may use the same device. For example, with a male adult preferring sports, a female adult preferring afternoon talk shows, and a three year old child preferring children's programming, the total information contained in the usage preference description 500 and the usage history description 502 will not be individually suitable for any particular user. The resulting composite data and its usage by the device is frustrating to the users because the device will not properly select and present audio and/or video content that is tailored to any particular user. To alleviate this limitation, the user preference description 20 may also include a user identification (user identifier) description 504. The user identification description 504 includes an identification of the particular user that is using the device. By incorporating a user identification description 504 more

than one user may use the device while maintaining a different or a unique set of data within the usage preference description 500 and the usage history description 502. Accordingly, the user identification description 504 associates the appropriate usage preference description(s) 500 and usage history description(s) 502 for the particular user identified by the user identification description 504. With multiple user identification descriptions 504, multiple entries within a single user identification description 504 identifying different users, and/or including the user identification description within the usage preference description 500 and/or usage history description 502 to provide the association therebetween, multiple users can readily use the same device while maintaining their individuality. Also, without the user identification description in the preferences and/or history, the user may more readily customize content anonymously. In addition, the user's user identification description 504 may be used to identify multiple different sets of usage preference descriptions 500 -- usage history descriptions 502, from which the user may select for present interaction with the device depending on usage conditions. The use of multiple user identification descriptions for the same user is useful when the user uses multiple different types of devices, such as a television, a home stereo, a business television, a hotel television, and a vehicle audio player, and maintains multiple different sets of preference descriptions. Further, the identification may likewise be used to identify groups of individuals, such as for example, a family. In addition, devices that are used on a temporary basis, such as those in hotel rooms or rental cars, the user identification requirements may be overridden by employing a temporary session user identification assigned by such devices. In applications where privacy concerns may be resolved or are otherwise

not a concern, the user identification description 504 may also contain demographic information of the user. In this manner, as the usage history description 502 increases during use over time, this demographic data and/or data regarding usage patterns may be made available to other sources. The data may be used for any purpose, such as for example, providing targeted advertising or programming on the device based on such data.

Referring to FIG. 23, periodically an agent 510 processes the usage history description(s) 502 for a particular user to "automatically" determine the particular user's preferences. In this manner, the user's usage preference description 500 is updated to reflect data stored in the usage history description 502. This processing by the agent 510 is preferably performed on a periodic basis so that during normal operation the usage history description 502 does not need to be processed, or otherwise queried, to determine the user's current browsing, filtering, searching, and device preferences. The usage preference description 500 is relatively compact and suitable for storage on a portable storage device, such as a smart card, for use by other devices as previously described.

Frequently, the user may be traveling away from home with his smart card containing his usage preference description 500. During such traveling the user will likely be browsing, filtering, searching, and setting device preferences of audio and/or video content on devices into which he provided his usage preference description 500. However, in some circumstances the audio and/or video content browsed, filtered, searched, and device preferences of the user may not be typically what he is normally interested in. In addition, for a single device the user may desire more than one profile depending on the season, such as football season, basketball season, baseball season, fall, winter, summer,

and spring. Accordingly, it may not be appropriate for the device to create a usage history description 502 and thereafter have the agent 510 "automatically" update the user's usage preference description 500. This will in effect corrupt the user's usage preference description 500. Accordingly, the device should include an option that disables the agent 510 from updating the usage preference description 500. Alternatively, the usage preference description 500 may include one or more fields or data structures that indicate whether or not the user desires the usage preference description 500 (or portions thereof) to be updated.

Referring to FIG. 24, the device may use the program descriptions provided by any suitable source describing the current and/or future audio and/or video content available from which a filtering agent 520 selects the appropriate content for the particular user(s). The content is selected based upon the usage preference description for a particular user identification(s) to determine a list of preferred audio and/or video programs.

As it may be observed, with a relatively compact user preference description 500 the user's preferences are readily movable to different devices, such as a personal video recorder, a TiVO player, a RePlay Networks player, a car audio player, or other audio and/or video appliance. Yet, the user preference description 500 may be updated in accordance with the user's browsing, filtering, searching, and device preferences.

Referring to FIG. 25, the usage preference description 500 preferably includes three different categories of descriptions, depending on the particular implementation. The preferred descriptions include (a) browsing preferences description 530, (b) filtering and search preferences description, 532 and (c) device preferences description 534. The browsing preferences

description 530 relates to the viewing preferences of audio and/or video programs. The filtering and search preferences description 532 relates to audio and/or video program level preferences. The program level preferences are not necessarily used at the same time as the (browsing) viewing preferences. For example, preferred programs can be determined as a result of filtering program descriptions according to user's filtering preferences. A particular preferred program may subsequently be viewed in accordance with user's browsing preferences. Accordingly, efficient implementation may be achieved if the browsing preferences description 530 is separate, at least logically, from the filtering and search preferences description 532. The device preferences description 534 relates to the preferences for setting up the device in relation to the type of content being presented, e.g. romance, drama, action, violence, evening, morning, day, weekend, weekday, and/or the available presentation devices. For example, presentation devices may include stereo sound, mono sound, surround sound, multiple potential displays, multiple different sets of audio speakers, AC-3, and Dolby Digital. It may likewise be observed that the device preferences description 534 is likewise separate, at least logically, from the browsing description 530 and filtering/search preferences description 532.

The browsing preferences description 530 contains descriptors that describe preferences of the user for browsing multimedia (audio and/or video) information. In the case of video, for example, the browsing preferences may include user's preference for continuous playback of the entire program versus visualizing a short summary of the program. Various summary types may be described in the program descriptions describing multiple different views of programs where these descriptions are utilized by the device to facilitate rapid non-linear browsing, viewing,

5 summary is preferred. In addition, browsing preferences
may also include descriptors describing parental control
settings. A switch descriptor (set by the user) should
also be included to specify whether or not the
preferences can be modified without consulting the user
10 first. This prevents inadvertent changing or updating of
the preferences by the device. In addition, it is
desirable that the browsing preferences are media content
dependent. For example, a user may prefer 15 minute
video highlight of a basketball game or may prefer to see
15 only the 3-point shots. The same user may prefer a
keyframe summary with two levels of hierarchy for home
videos.

The filtering and search preferences description 532 preferably has four descriptions defined therein, depending on the particular embodiment. The keyword preferences description 540 is used to specify favorite topics that may not be captured in the title, category, etc., information. This permits the acceptance of a query for matching entries in any of the available data fields. The content preferences description 542 is used to facilitate capturing, for instance, favorite actors, directors. The creation preferences description 544 is used to specify capturing, for instance, titles of favorite shows. The classification preferences description 546 is used to specify descriptions, for instance, a favorite program category. A switch descriptor, activated by the user, may be included to specify whether or not the preferences may be modified without consulting the user, as previously described.

35 The device preferences description 534 contains
descriptors describing preferred audio and/or video
rendering settings, such as volume, balance, bass,

treble, brightness, contrast, closed captioning, AC-3, Dolby digital, which display device of several, type of display device, etc. The settings of the device relate to how the user browses and consumes the audio and/or video content. It is desirable to be able to specify the device setting preferences in a media type and content-dependent manner. For example the preferred volume settings for an action movie may be higher than a drama, or the preferred settings of bass for classical music and rock music may be different. A switch descriptor, activated by the user, may be included to specify whether or not the preferences may be modified without consulting the user, as previously described.

Referring to FIG. 26, the usage preferences description may be used in cooperation with an MPEG-7 compliant data stream and/or device. MPEG-7 descriptions are described in ISO/IEC JTC1/SC29/WG11 "MPEG-7 Media/Meta DSs (V0.2)", August 1999, incorporated by reference herein. It is preferable that media content descriptions are consistent with descriptions of preferences of users consuming the media. Consistency can be achieved by using common descriptors in media and user preference descriptions or by specifying a correspondence between user preferences and media descriptors. Browsing preferences descriptions are preferably consistent with media descriptions describing different views and summaries of the media. The content preferences description 542 is preferably consistent with, e.g., a subset of the content description of the media 553 specified in MPEG-7 by content description scheme. The classification preferences description 544 is preferably consistent with, e.g., a subset of the classification description 554 defined in MPEG-7 as classification description scheme. The creation preferences description 546 is preferably consistent with, e.g., a subset of the creation description 556 specified in MPEG-7 by creation description scheme. The keyword preferences description

540 is preferably a string supporting multiple languages and consistent with corresponding media content description schemes. Consistency between media and user preference descriptions is depicted or shown in FIG. 26 by double arrows in the case of content, creation, and classification preferences.

Referring to FIG. 27, the usage history description 502 preferably includes three different categories of descriptions, depending on the particular implementation. The preferred descriptions include (a) browsing history description 560, (b) filtering and search history description 562, and (c) device usage history description 564, as previously described in relation to the usage preference description 500. The filtering and search history description 562 preferably has four descriptions defined therein, depending on the particular embodiment, namely, a keyword usage history description 566, a content usage history description 568, a creation preferences description 570, and a classification usage history description 572, as previously described with respect to the preferences. The usage history description 502 may contain additional descriptors therein (or description if desired) that describe the time and/or time duration of information contained therein. The time refers to the duration of consuming a particular audio and/or video program. The duration of time that a particular program has been viewed provides information that may be used to determine user preferences. For example, if a user only watches a show for 5 minutes then it may not be a suitable preference for inclusion the usage preference description 500. In addition, the present inventors came to the realization that an even more accurate measure of the user's preference of a particular audio and/or video program is the time viewed in light of the total duration of the program. This accounts for the relative viewing duration of a program. For example watching 30 minutes

of a 4 hour show may be of less relevance than watching 30 minutes of a 30 minute show to determine preference data for inclusion in the usage preference description 500.

5 Referring to FIG. 28, an exemplary example of an audio and/or video program receiver with persistent storage is illustrated. As shown, audio/video program descriptions are available from the broadcast or other source, such as a telephone line. The user preference
10 description facilitate personalization of the browsing, filtering and search, and device settings. In this embodiment, the user preferences are stored at the user's terminal with provision for transporting it to other systems, for example via a smart card. Alternatively,
15 the user preferences may be stored in a server and the content adaptation can be performed according to user descriptions at the server and then the preferred content is transmitted to the user. The user may directly provide the user preferences, if desired. The user
20 preferences and/or user history may likewise be provided to a service provider. The system may employ an application that records user's usage history in the form of usage history description, as previously defined. The usage history description is then utilized by another
25 application, e.g., a smart agent, to automatically map usage history to user preferences.

Additional Attributes and Descriptors

In The Description and The Description Scheme

30 The present inventors came to the realization that additional functionality for the system may be achieved by the incorporation of particular types of information in the descriptions and description schemes. A description scheme is a data model of descriptions. It
35 specifies the descriptors and their syntax as they are used in the description. In what follows, use the terms description and description scheme may be used

interchangeably since they both correspond to describing media and user preferences. An explanation of the additional attributes and descriptors in the descriptions will be provided, followed by an example of portions of example descriptions.

After further consideration, there is a need for many users to maintain multiple separate user preference descriptions. Multiple user preference descriptions may correspond to, for example, different locations (e.g., at home, at the office, away from home, stationary versus traveling in a vehicle), different situations, different times (e.g., different days, different seasons), different emotional states of the user (e.g., happy mood versus tired or sad), and/or persistence (e.g., temporary usage versus permanent usage). Further, the user preference descriptions may include differentiation for different terminals with different primary functionalities (e.g., a personal video recorder versus a cell phone). In addition, available communication channel bandwidth at different locations or situations may use different preferences. Also, the preference of a user for the length of an audiovisual summary of a video program for downloading may be different. The user in different usage conditions may use the user identification description scheme as a basis to distinguish between different devices and/or services. An example of different conditions may include a television broadcast receiver and a cellular telephone.

In addition to maintaining multiple user preferences for a particular user based on the aforementioned conditions, the present inventors also came to the realization that the different locations, different situations, different emotional states, different seasons, and/or different terminals (etc.), may likewise be used as the basis for distinguishing between the user preference descriptions.

One technique to permit a particular user to have multiple preference descriptions and distinguishing them from one another is by using different usernames or by using a versioning mechanism, such as a version descriptor in the identification description scheme, as described later.

As previously described, the system may include multiple user preference descriptions for a particular user. With multiple descriptions, the system may express the different user preferences with different granularity, e.g., a greater or lesser amount of detail. The increased granularity (sparseness) may be merely the result of applying a filter to the user preference description that further reduces the amount of data. In other words, the structure of the usage preference description may be identical with the difference being the result of the filter further reducing the data. In another embodiment, the variable granularity results in a different size of the data contained in the user preferences, which may be based upon, if desired, the location and/or application of the user. User preferences with increased granularity may be especially suitable for storage on portable memory devices with limited memory capability. Likewise, the granularity may be applied to the usage history.

Another aspect of the present invention permits the user preferences (and history) to be based upon the media type, media source, or content (e.g., music versus video, radio versus television broadcast, and/or sports video versus home video). These preferences relate to the audio and/or video itself, as opposed to a third party characterization of the desirability of the multimedia. The inclusion of this information permits a reduction in the computational processing requirements depending on the media type, media source, and/or content of the media.

Another feature that may be included in the system is a protection attribute for each, or a selected set of, component of the user descriptions. The protection attributes specifies the access right of a system or service provider, typically a party other than the user himself, to the user's descriptions or any component thereof. In one embodiment, the protection attributes may be specified by a binary value that indicates the user's desire to permit others access to such data. One technique to implement the protection attribute is to include a protection attribute as a primitive attribute that is contained by all relevant parts of the user description scheme.

Descriptors and description schemes for browsing preferences may be aligned with particular types of multimedia summary description schemes that are contained in ISO/IEC JTC1/SC29/WG11 N3246, "MPEG-7 Generic AV Description Schemes, Working Draft v2.0", Noordwijkerhout, March 2000. This allows the user to specify the type of a particular visual summary of an audiovisual program, and the duration of a summary that is in the form of a visual highlight. However, after further consideration the present inventors have determined that specification of the preferred minimum and maximum amount of data permitted in an audiovisual summary significantly enhances the system capability. Such a provision provides, for example, the capability of the user effectively browsing audiovisual summaries of content over channels with limited bandwidth and using terminals with different limitations. With a terminal connected to a bandwidth limited channel, the user may specify preference for a relatively short highlight of the program, while with a terminal that is connected to a higher bandwidth channel, the user may specify preference for a longer highlight of the program. Such a set of channels may be mobile channels and cable channels. In addition, for terminals that are not capable of

displaying frames at a video rate, the user may prefer keyframe summaries consisting of a maximum number of keyframes appropriate for the communication channel bandwidth. To achieve these enhancements, the present
5 inventors propose using descriptors in the browsing preferences description (and description scheme, or other preferences description) specifying the minimum, maximum, and exact number of keyframes, and minimum, maximum, and exact duration of audio and/or visual highlights.

10 As described, the description scheme is adaptable to express the preferred minimum and maximum amount of visual material to adapt to different viewing preferences as well as terminal and communication channel bandwidth limitations. This implementation may be
15 achieved by the following descriptors included in the browsing preferences description scheme:
MaxNumOfKeyframes, MinNumOfKeyframes, NumOfKeyframes, MaxSummaryDuration, MinSummaryDuration, and
SummaryDuration. The MaxNumOfKeyframes and
20 MinNumOfKeyframes preference descriptors specify, respectively, the maximum and minimum number of keyframes in the keyframe-summary of a video program. Depending on the known bandwidth conditions of a known connection that the user uses regularly, he or she may specify these
25 descriptors. The MaxSummaryDuration and MinSummaryDuration descriptors specify, respectively, the maximum and minimum temporal duration of an audiovisual highlight summary. Again, depending on user's taste, terminal, and channel limitations, the user may specify
30 these descriptors. The MaxSummaryDuration and MinSummaryDuration descriptors apply to preferences for audio signals as well as where audio highlights may have been generated by audio skimming methods. User's
35 browsing preference descriptions may be correlated with media descriptions by a filtering agent 520 in Fig. 24 in order to determine media descriptions that contain summary descriptions that match user's preference

descriptions and provide the user the associated summarized media in the preferred type of summary.

An additional descriptor that may be introduced is an abstraction fidelity descriptor for universal multimedia access application, where fidelity of a summary abstraction of a program is described. This can correspond to the variation fidelity descriptor defined in ISO/IEC JTC1/SC29 WG11 N3246, "MPEG-7 Multimedia Description Schemes, Working Draft v2.0", Noordwijkerhout, March 2000. This provides an alternative to the explicit specification of the duration and bounds on the number of keyframes. A Segment Theme descriptor(s) may describe the preferred theme, or point of view, of a segment, e.g., a video or audio clip, annotated with its theme or emphasis point. For example, the theme may specify characteristics of the content of the theme. Such characterization may include a goal from your favorite team, 3-point shots from your favorite player, etc. Specifying these descriptor(s) and also ranking them enables a client application or a server to provide to the user segments according to preferred themes (and/or their ranking) matching to the their labels or descriptors at the segment level, or provide users with pre-assembled highlights composed of segments with labels matching the SegmentTheme preference.

Existing filtering and search user preference descriptions are directed to techniques of using the audiovisual content in an effective manner by finding, selecting and consuming the desired audiovisual material, while focusing on the content of the audiovisual materials. While such descriptions are beneficial, the present inventors came to the further realization that the identification of the source of the material, in contrast to merely its content, provides beneficial information for the processing and presentation of the audiovisual materials. For example, the source of the content may be from terrestrial sources, digital video

disc, cable television, analog broadcast television, digital broadcast television, analog radio broadcasts, and digital radio broadcasts. The inclusion of this information permits the user to select among these

5 different sources and increase effectiveness by narrowing down the choices to those sources that are available to the user, such as terrestrial broadcast which is more widely available than satellite broadcast. For example, user may describe user's preference for "Star Trek"

10 episodes that are available from terrestrial broadcast channels only.

This source distinction and identification may be performed by including a source preferences description scheme under the filtering and search

15 preferences description scheme (or other description scheme). Accordingly, the search and preferences description scheme may include from zero or one (or more if desired) source preferences description scheme. The source preferences description scheme may be derived from

20 the Media Format description scheme or Publication Description Scheme specified in ISO/IEC JTC1/SC29/WG11 N3247, MPEG-7 Multimedia Description Schemes, Experimentation Model (v2.0) Noordwijkerhout, March 2000.

25 Another feature that may be included in the system, in addition to the user's preferences, is the user's negative preferences. The negative preferences may include the user's dislikes and their relative rankings. By specifying the negative preferences, the system is less likely to select such matching preferences.

30 This may be implemented, for example, by permitting positive and negative values to the preferencevalue descriptor.

Another feature that may be included in the system is the specification of the user's preferences as a relative preference measure of a particular set of user preferences with respect to another set of preferences, such as for example, by using

35 BetterThan and WorseThan descriptors. This permits an implicit relative ranking of preferences even in the absence of a preference value descriptor for each preference

002010"30611560

set. This may be implemented, for example, by including Betterthan and WorseThan descriptors in the filtering and search preferences descriptions.

Expression of the Additional Attributes

The following descriptions are expressed in XML (Extensible Markup Language), incorporated by reference herein. It is to be understood that any other description language may likewise be used.

The definition of the user preference description may be as follows.

```
<UserPreference>
  <UserIdentifier protection="true" userName="paul"/>
  <UsagePreferences allowAutomaticUpdate="false">
    <BrowsingPreferences>
      ...
    </BrowsingPreferences>
    <FilteringAndSearchPreferences>
      ...
    </FilteringAndSearchPreferences>
    <DevicePreferences>
      ...
    </DevicePreferences>
  </UsageHistory>
  ...
</UsageHistory>
</UserPreference>
```

The primitive attributes "protection" and "allowAutomaticUpdate" may be instantiated in the UserIdentifier, Usage Preferences, and Usage History descriptions and all its relevant parts, namely, in Browsing Preferences description, Filtering and Search Preferences description, Device Preferences description, and sub-description schemes of the Usage History description Scheme.

The "allowAutomaticUpdate" attribute (set by the user) should be included in a description scheme specifying whether or not the preferences can be automatically modified (e.g., by an agent utilizing the usage history description) without consulting with the user.

The protection attribute should be included in a description specifying whether the user allows the system to make preference/history public or not. When the user agrees to make some parts of his preference/history public, for example, to service providers, the service providers can collect this information and then serve to the user contents that are tailored to the user's history/preferences. In the above example description, the user prefers to keep his username private. He also does not wish the system to automatically update his preferences.

The user identification description serves the purpose of an identifier that distinguishes a particular instantiation of the user description scheme from other instantiations for other users or other instantiations for the same user for different usage conditions and situations.

The username descriptor may identify a specific user from other users. In a home setting, each member of the household may be identified using a username that is unique in the household for all devices that the members of that household use on a regular basis. A username can also be used to distinguish the user description scheme of not only an individual but also a group of people, e.g., the family. Those devices that are used on a temporary basis, potentially by many different people, (such as those in hotel rooms or rental cars) may assign temporary session identifications to ensure uniqueness of identifications.

Alternatively, a version descriptor may also be included in the user identifier description to define different versions of the user descriptions (preferences and usage history) associated with a particular username. Through the mechanism of the version, a person can specify different preferences and usage history, corresponding to different locations (at home, at the office, away from home, stationary versus traveling in a vehicle), different situations, different emotional states (happy versus sad), different seasons, etc. Different user descriptions are distinguished by distinct version descriptors. The type of the version descriptor, may be for example, an integer, a string, or expressed as an attribute of the user identification description scheme.

The usage preference description may include a PreferenceType description, distinguishing a particular set of preferences or history according to time,

or place, or a place and time combination. The definition of the usage preference description may be as shown in the following example, where place is "office" and time period is "8 hours starting from 8 AM"

```

5         <PreferenceType>
           <Place>
             <PlaceName xml:lang="en">Office</PlaceName>
           </Place>
           <Time>
10          <TimePoint>
             <h>8</h>
           </TimePoint>
           <Duration>
             <No_h>8</No_h>
15          </Duration>
           </Time>
        </PreferenceType>

```

20 The preferencetype descriptor may be used to identify the preference type of one or more set of preferences. As previously described, a user may have different preferences depending on the user's situation, location, time, season, and so on.

25 The browsing preferences description may describe preferences of the user for browsing multimedia information. In essence, this description expresses the user's preferences for consuming (viewing, listening) a multimedia information. This browsing preferences description may include for example, a Summary Preferences description. The browsing preferences description may include in the case of video, for example, the user's preferences for continuous playback of the entire program

30 versus visualizing a short summary of the program. Various summary types are specified in the Summary Description Scheme in ISO/IEC JTC1/SC29 WG11 N3246, "MPEG-7 Multimedia Description Schemes, Working Draft v2.0", Noordwijkerhout, March 2000, including a keyframe summary, a highlight summary, etc., where

parameters of the various summary types may also be specified by summary descriptions, e.g., the time duration of the video highlight summary.

The browsing preferences description scheme may include one or more of the following non-exhaustive list of descriptors and descriptions in its description scheme.

(A) The minimum number of keyframes (MinNumOfKeyframes) and the maximum number of keyframes (MaxNumOfKeyframes) descriptors may be included. These descriptors specify the user's preference for minimum and maximum number of frames in a keyframe summary of an audiovisual program. A user can specify these descriptors according to personal taste, situation, etc., and according to channel bandwidth and terminal resource limitation.

(B) The minimum duration (MinSummaryDuration) and the maximum duration (MaxSummaryDuration) descriptors may be included. These descriptors specify the user's preference for the length of a highlight summary composed of key clips in the video. These descriptors may also, for example, be applied to an audio-only material. A user can specify these descriptors according to personal taste, situation, etc., and according to channel bandwidth and terminal resource limitations.

An example for Summary Preferences description that can be included in usage preferences description is provided below.

```
</UsagePreferences>
</BrowsingPreferences>
<SummaryPreferences>
  <SummaryTypePreference>keyVideoClips</SummaryTypePreference>
  <MinSummaryDuration><m>3</m><s>20</s></MinSummaryDuration>
  <MaxSummaryDuration><m>6</m><s>40</s></MaxSummaryDuration>
</SummaryPreferences>
</BrowsingPreferences>
</UsagePreferences>
```

(C) The abstraction fidelity descriptor for universal multimedia access application relates to fidelity of a summary abstraction of a program. This preference descriptor may correspond to the variation fidelity descriptor contained in the media's variation description specified by Variation Description Scheme in ISO/IEC JTC1/SC29 WG11 N3246, "MPEG-7 Multimedia Description Schemes, Working Draft v2.0", Noordwijkerhout, March 2000. Alternatively, the duration and number of keyframes may be defined as the fidelity descriptor.

(D) The SegmentTheme descriptor(s) may be included, which describes the theme or point of view of a segment, e.g., a video or audio clip annotated with its theme or emphasis point. An example summary preference description expressing preference for video segments (clips) labeled as "Goal from Spain" and "Replay of Goal from Spain" is as follows:

```

...
</UsagePreferences>
  </BrowsingPreferences>
    <SummaryPreferences>
      <SummaryTypePreference>KeyVideoClips</SummaryTypePreference>
      <SegmentTheme>Goal from Spain</SegmentTheme>
      <SegmentTheme>Replay of goal from Spain</SegmentTheme>
    </SummaryPreferences>
  </BrowsingPreferences>
</UsagePreferences>

```

(E) The frame frequency value descriptor may be included to specify the temporal sampling frequency of video frames that can be visualized in the browser. The frames provide a visual summary. Depending on the browser, they may also provide clickable entry points to the video. The user may click and start playing back the video starting from that frame. The frame frequency value descriptor provides similar functionality in terms of shots of the video.

The source preference description describes the preferred source of multimedia information, such as the broadcast or storage medium type (e.g., terrestrial, satellite, DVD), broadcast channel identifier, etc. An example user preference description expressing preference for Star Trek episodes available from terrestrial broadcast is as follows.

```

<UserIdentifier protection="true" userName="paul"/>
<UsagePreferences allowAutomaticUpdate="false">
  <FilteringAndSearchPreferences protection="true">
    <PreferenceValue>5</PreferenceValue>
    <CreationPreferences>
      <Title xml:lang="en" type="original">Star Trek</Title>
      </CreationPreferences>
      <SourcePreferences>
        <PublicationType>Terrestrial Broadcast</PublicationType>
      </SourcePreferences>
    </FilteringAndSearchPreferences>
  </UsagePreferences>
</UserIdentifier>

```

The filtering and search preferences description includes at least one of the descriptors of preferred program title, genre, language, actor, creator of the program. An example description where user's preference is for news programs in English is given below. Such description may be included in user's smart card when he travels to Japan, for example. Note that this particular preference description is identified as being specific to Japan and differentiated by choosing an appropriate user name.

```

<UserIdentifier protection="true" userName="paul_in_Japan"/>
<UsagePreferences allowAutomaticUpdate="false">
  <FilteringAndSearchPreferences protection="true">
    <PreferenceValue>100</PreferenceValue>
    <ClassificationPreferences>
      <Language>
        <LanguageCode>en</LanguageCode>
      </Language>
      <Genre>News</Genre>

```

```
</ClassificationPreferences>  
</FilteringAndSearchPreferences>  
</UsagePreferences>  
</UserIdentifier>
```

5

The PreferenceValue descriptor provides a technique for prioritizing filtering and search preferences, such as the value indicating the degree of user's preference or non-preference. Non-preferences may be expressed by assigning a negative (opposite) value to the preference value descriptor.

10

The betterthan and worsethan descriptors may describe which instantiation of preferences the user likes or dislikes relatively more compared to another instantiation, where different instantiations are identified using the filtering and search preference type descriptor. This provides robustness against changes in the preference value descriptor automatically, for example, by an agent.

15

The filtering and search preferences description may also contain a description of a preferred review to express user's desire for searching for programs that are favorably reviewed by specific individuals. For example, preference for movies reviewed by movie critics Siskel and Ebert and found to be "two-thumbs- up" may be described and included in the filtering and search preferences description.

20

An overview of the entire description scheme is shown in FIG. 29.

25

The terms and expressions that have been employed in the foregoing specification are used as terms of description and not of limitation, and there is no intention, in the use of such terms and expressions, of excluding equivalents of the features shown and described or portions thereof, it being recognized that the scope of the invention is defined and limited only by the claims that follow.